Attractors of entrepreneurial activity: Universities, regions, and alumni entrepreneurs

Anders Broström, Apostolos Baltzopoulos

Attractors of entrepreneurial activity: Universities, regions, and alumni entrepreneurs

Journal:	*Regional Studies*
Manuscript ID:	CRES-2010-0333.R1
Manuscript Type:	Main Section
JEL codes:	R30 - General < R3 - Production Analysis and Firm Location < R - Urban, Rural, and Regional Economics, I28 - Government Policy < I2 - Education < I - Health, Education, and Welfare, J61 - Geographic Labor Mobility|Immigrant Workers < J6 - Mobility, Unemployment, and Vacancies < J - Labor and Demographic Economics, O18 - Regional, Urban, and Rural Analyses < O1 - Economic Development < O - Economic Development, Technological Change, and Growth
Keywords:	university entrepreneurship, regional impact, location choice analysis

Attractors of entrepreneurial activity

Universities, regions, and alumni entrepreneurs

APOSTOLOS BALTZOPOULOS*† and ANDERS BROSTRÖM†‡

*Nordregio, Nordic Centre for Spatial Development, Box 1658, SE-111 86, Stockholm, Sweden, Email: apostolos.baltzopoulos@nordregio.se

†Division of Economics, Royal Institute of Technology, SE-100 44, Stockholm, Sweden, Emails: apostolos.baltzopoulos@abe.kth.se and anders.brostrom@abe.kth.se

‡CESIS, Centre of Excellence for Science and Innovation Studies, Drottning Kristinas väg 30, SE-100 44, Stockholm, Sweden

(Received October 2010: in revised form June 2011)

Abstract

This paper investigates how universities may affect regional entrepreneurship through the localisation decisions of entrepreneurial alumni. Empirically a comprehensive, individual-level dataset from Sweden for the period 2003-2005 is employed. The results suggest that even when controlling for their spatial history, individuals have an increased propensity to set up in the region where they studied. This effect is found to substitute for both urbanisation economies and localisation economies as drivers of regional-level entrepreneurship. Thus, the present analysis provides evidence on how universities affect regional economic development that complements the strong focus on spin-off activities by university researchers in previous studies.

Keywords - university entrepreneurship, regional impact, location choice analysis

JEL classifications - I23, L26, O18, R30

INTRODUCTION

The importance of universities[1] for regional economic development has been analysed from different economic perspectives (ANDERSSON et al., 1990; FELSENSTEIN, 1996, 1996; PHELPS, 1998; CHESHIRE and MALECKI, 2004). Alongside direct spending effects (FLORAX, 1992), universities are associated with productivity gains and innovation in existing firms, with effects on new firm-formation and industry location choices and therefore with long-term regional growth (BANIA et al., 1993; GOLDSTEIN and

RENAULT, 2004). A common caveat is that while such mechanisms seem to enhance regional economic growth at some locations, this effect is mediated by different regional characteristics such as the density of population and business activity (VARGA, 1998; GOLDSTEIN and DRUCKER, 2006) and the region's industrial structure (BRAUNERHJELM, 2008). However, the mechanisms through which universities affect regional growth and the specific reasons for the heterogeneous impact of universities across regions have not yet been pinned down. There exists, nonetheless, a prevailing expectation that the process of entrepreneurship has a considerable role to play in this context. This study investigates how universities may affect regional entrepreneurship through the localisation decisions of entrepreneurial alumni.

Despite considerable policy interest in the role of higher education institutions for regional entrepreneurship (FELDMAN, 2001; ETZKOWITZ and KLOFSTEN, 2005), a large majority of studies of the university-entrepreneurship linkage focus on the case of academic entrepreneurship, i.e. on academic researchers' engagement in start-up ventures (LINDHOLM-DAHLSTRAND, 1997; for a review, see ROTHERMAEL et al., 2007). Non-faculty entrepreneurship activities have, with the exception of occasional studies from singular universities (e.g. HSU et al., 2007), only been examined insofar as it has taken a path over university-owned science parks and incubators (HISRICH and SMILOR, 1988; for a review, see LINK and SCOTT, 2007). In addressing the phenomenon of alumni entrepreneurship, this study examines a mechanism which may explain more of the

measured impact of universities on a region than what has hitherto been acknowledged (GERTLER, 2010).

The assumption that there may be such a connection is based on entrepreneurship theory, in which entrepreneurial activity is considered to be a truly regional phenomenon (STERNBERG and WENNEKERS, 2007). On the one hand, (necessity-based) entrepreneurship may be a response to the desire of the individual to live in a certain region, and a failure to find a suitable job in existing firms. On the other, (opportunity-based) entrepreneurship may arise as a consequence of the recognition of opportunities in markets with which the nascent entrepreneur is familiar. For both types of entrepreneurship, the location choice is likely to be conditioned by personal networks. Familiarity with a region and its different markets as well as personal contacts is particularly valuable in the establishment-phase of a new venture (STAM, 2007). Therefore, this paper tests the hypothesis that higher education institutions may bolster entrepreneurship in a region simply by pulling talented people to the region, where they may then choose to remain – possibly in a role as an entrepreneur. Furthermore the analysis considers the relationship between this "alumni effect" on the location choices of entrepreneurs and the theoretical concept of agglomeration economies. In examining the alumni effect as a possible substitute for the traditional forces of urbanisation and localisation economies, respectively, two questions are asked. Does a university affect regional entrepreneurship the most in urban or non-urban regions? Does alumni

entrepreneurship strengthen industrial clustering effects, or does it provide a means for diversification of the regional economy?

This study is one of the first to explore differences in entrepreneurship across regions utilising individual-level data, following the path pioneered by EVANS and LEIGHTON (1989). This approach allows dealing with the problem of the counterfactual – a typical problem for all kinds of impact studies, which is particularly difficult to solve in studies of the impact of universities (SIEGFRIED et al., 2007) – in a satisfactory way. Results based on a comprehensive Swedish database support the hypothesis that entrepreneurs will exhibit an increased propensity to start their firm in their region of studies. This tendency appears to be stronger in more peripheral areas of Sweden that in the three major urban centres. The analysis also indicates that the pull effect of universities substitutes rather than complements that of localization externalities. Together, these findings suggest that universities, through the mechanism of alumni entrepreneurship, play a particularly interesting role for the renewal of non-urban regional economies.

The rest of the paper is organized in the following manner: Section 2 discusses the influence of universities on regional entrepreneurship and presents the hypotheses addressed while section 3 presents the approach applied to empirically test them. Section 4 presents and discusses the results of the econometric analysis and section 5 summarizes and concludes the paper.

THE INFLUENCE OF UNIVERSITIES ON REGIONAL ENTREPRENEURSHIP

As noted by DRUCKER and GOLDSTEIN (2007), the role of universities for labour mobility has not been well-researched. HUFFMAN and QUIGLEY (2002) study 950 graduates of two California universities, and their propensity to stay in the state after graduation. Similar studies, both covering single U.S. universities, are presented by BLACKWELL at al. (2002) and FELSENSTEIN (1995). Case-study research of particular universities or regions has also suggested that universities may play a significant regional role as attractors of talent (SAXENIAN and HSU (2001). GROEN (2004) uses data on 30 selective colleges and universities and reports only a modest link between attending college in a state and working in the state. HOARE and CORVER (2010) study the mobility patterns of young UK higher education graduates into their first employment across 12 regions. They find that prior local experience from residence before studies and from studies – and in particular a combination of these types of local ties – are associated with greater tendency to local recruitment for almost all regions.

The assumption underlying the present research is that higher education experiences increase the attractiveness of the location where studies took place in future location decisions. Higher education experiences can constitute a foundation for personal networks and, to the extent that these networks are localised, thereby create localised social capital that may be utilised in entrepreneurial activities (STUART and SORENSEN, 2003). In

particular, this applies to young individuals who choose to relocate in order to attend a specific university and find themselves in a novel environment, "cut-off" from friends and family, and subsequently proceed with building an entirely new social network. This network may well exceed the strict confines of the halls of a university but is still centred in their new place of residence. The importance of social networks for entrepreneurship is widely acknowledged. GREVE and SALAFF (2003), for example, stress how entrepreneurs require information, capital, skills, and labour to start their business activities and while they hold some of these resources themselves, they often complement their resources by accessing their contacts (ALDRICH and ZIMMER, 1986; COOPER et al., 1995; HANSEN, 1995, JOHANNISSON, 1988). Beyond social networks, higher education experiences can also be expected to increase the alumni's knowledge about the local market. These arguments are of course not unique to higher education experiences; the entire spatial history of an entrepreneur can be expected to influence his or her localisation choice when establishing a new firm by affecting his or her embededdness in different regions (STUART and SORENSEN, 2003). In particular, the entrepreneur can be expected to be biased towards setting up in his or her place of birth and in regions where he or she has previous working experience. However, it is postulated that the networks that individuals develop in the formative years of their university studies are important enough to affect future business venturing, both directly (direct utilisation of university-based networks) and indirectly (i.e. that the individual´s choice of location for working may be influenced by her or his choice of location for studies). Strengthening these effects, personal ties to a specific region created during studies (e.g. students forming a family during their studies

which) may introduce significant switching costs to the alternative of relocating prior to business founding. These assumptions lead to the following hypothesis:

H1: Controlling for their region of birth and their recent employment history start-up founders have an increased propensity to set up in the region where they studied

Considering the heterogeneous impact of universities on economically well-developed and less-developed regions that has been identified by previous studies, an important extension to the analysis is to examine to what extent such differences can be explained by university-induced entrepreneurship. The main question is how the effect described in Hypothesis 1 – hereafter the *alumni effect* – interacts with the traditional pull-factors of localisation analysis. In particular: is the alumni effect best understood as a substitute or a complement to the forces of agglomeration economies?

The theory of agglomeration economies posits that firms benefit from localised dynamics of sharing, matching and learning mechanisms (DURANTON and PUGA, 2004). Considering the question of their relation to the alumni effect, the existing literature offers no univocal theoretical prediction to build on. On the one hand, successful sharing, matching and learning activities could be considered to be facilitated by a firm's local embeddeddness, suggesting the alumni effect to be stronger in regions with strong agglomeration economies. On the other hand, following WEBBER (1972), the reduced uncertainty associated with embeddedness can be considered to reduce the need for

external economies of agglomeration. In view of these contradictory theoretical predictions, a more detailed discussion about localization and urbanization economies follows.

The impact of agglomeration on firm decisions is typically considered to be twofold (JOHANSSON, 2004), a division that goes back to OHLIN (1933). First, large agglomerations are seen as attractive milieus for entrepreneurial ventures due to existence of a rich and diverse labour pool, a big local consumer market, and well developed infrastructure, including public facilities and an abundance of secondary services. These effects, which accrue across sectors, are usually labelled as urbanisation economies. Secondly, firms are believed to be able to benefit from co-location with other firms in related sectors through the presence of localisation economies. Following MARSHALL (1920) and MYRDAL (1957), localisation economies have been considered to arise from the agglomeration of firms, which allows suppliers to specialise, increases the supply of specialised labour and spurs local demand.

Recently, the nature of localisation economies has been re-examined and their role for entrepreneurship re-interpreted. Traditionally, following the new economic geography theories of e.g. KRUGMAN (1991), entrepreneurs have been assumed to be affected by the forces of agglomeration when making location choices (KEEBLE and WALKER, 1994; BOSCHMA and WENTING, 2007). However, ROSENTHAL and STRANGE (2001) found that agglomeration economies are mainly benefiting co-localised firms through labour

market pooling benefits, which would suggest that agglomeration economies are more important for the growth of existing firms than for the success of very new firms. This interpretation can also be supported by the findings of DURANTON and PUGA (2000, 2001) who find that while new plants tend to be attracted to set up in diversified areas, relocations are more likely to go to specialised areas. Recently, SORENSON and AUDIA (2000), KLEPPER (2007), BUENSTORF and GUENTHER (2007) and BUENSTORF and KLEPPER (2009) have suggested that localisation economies should not be understood as a phenomenon that gives any new firm an incentive to locate close to firms in related industries. Industrial clustering can in this interpretation rather be understood as the observed outcome of localised processes of heritage. This explanation for the clustering of industries emphasises spin-off dynamics from existing firms in combination with a tendency to locate new firms close to where the entrepreneur lived and worked at the time of firm start-up.

This evidence seems to suggest that the kinds of networks and experiences that are accumulated in higher education would not increase the likelihood that an alumni entrepreneur chooses to start a firm in the specific kinds of industries that are already clustered in the region. Rather on the contrary, university-induced networks could provide an alternative to networks accrued in working life, and entrepreneurial opportunities that are recognised during studies could provide an alternative to entrepreneurial opportunities that are recognised when working in a region. The following hypothesis is formulated:

H2: When localisation decisions are driven by an alumni effect, the impact of localisation economies decreases

The next issue considered is how the pull effect of universities interacts with urbanisation economies in the location decisions of alumni entrepreneurs. Urban regions have advantages for new firms that are particularly well articulated for the type of knowledge-based service firms which are the hallmark of entrepreneurs with higher education (WAGNER and STERNBERG, 2004; LEE et al., 2004). There are therefore reasons to expect that the alumni effect will have a stronger impact on entrepreneurial alumni from urban regions than on other alumni; entrepreneurial impulses and networks created during education may not be sufficient to encourage local entrepreneurship if the region does not offer sufficient market size and advanced demand. However, there are also plausible arguments for a reversed relationship. From the point of view of the alumni of non-urban regions, self-employment may act as a substitute for the richer labour market opportunities of urban regions. Entrepreneurship may thus become attractive for those with strong preferences against leaving their (non-urban) region. This view is supported by FIGUEIREDO et al. (2002) who find that urbanization economies are only a significant factor for the localization decisions of entrepreneurs who move from their previous location, and not for entrepreneurs who stay on. Interestingly, Figuerido et al. present the only evidence on this issue that similar to the current analysis explores the level of individual decisions rather than more aggregate regional measurements. In lack of more

in-depth guidance on this issue, it is proposed that the total effect goes towards strengthening the effect of universities on entrepreneurship in non-urban regions.

H3: When localisation decisions are driven by an alumni effect, the impact of urbanisation economies decreases

When controlling for other characteristics of a region that may affect the location choices of entrepreneurs, it is interesting to note that when modelling this choice on the individual level, several conceptually separate factors can be grouped together under a general control for regional heterogeneity. Among such factors, that can be considered equally important for all types of entrepreneurs, are heterogeneous incentive structures for entrepreneurship across regions such as tax structures (WAGNER and STERNBERG, 2004), characteristics of the local population and institutional factors characterising the regional economic milieu (ARMINGTON and ACS, 2002). The notion that several regional characteristics may be meaningfully grouped together in location choice analysis is supported by the striking finding of cross-regional differences in entrepreneurship as a highly persistent phenomenon over time. Evidence for such cross-regional patterns is convincing for several countries (ACS and ARMINGTON, 2004, US; FRITSCH and MUELLER, 2007, Germany; ANDERSSON and KOSTER, 2010, Sweden).

METHODOLOGY

Data

The above hypotheses are empirically tested using census data from Sweden that describe the country's business and employment dynamics over the period 1985 to 2005. We use a comprehensive micro-database covering all working individuals in the Swedish economy. All individuals are matched to the firms and establishments they own or are employed in. Our database furthermore has three very useful characteristics. First of all, firm dynamics are monitored and reported in great detail. This allows us to identify whether a new firm has been the result of a split or a merger of previously existing firms or whether it is a Greenfield start-up. In this paper an entrepreneur is defined as the owner of a newly-found start-up. Secondly, the database includes information on the geographic location of the place of residence, work, and birth of each individual at the municipality level, allowing for extremely detailed labour mobility analyses. Thirdly, the level and type of education of all individuals is recorded. In particular for people that received a university education in Sweden, the year and region of graduation are reported without fail. Earlier studies on alumni have been forced to use data created through cross-sectional surveys, since universities themselves rarely has been able to provide comprehensive follow-up records of the career paths of alumni. By contrast, our database allows tracking the entire employment/self-employment history of all people graduating from a Swedish university (as long as they joined the Swedish job market).

The main goal is to test whether an individual entrepreneur will show an increased propensity to choose to set up his firm in the region where he or she attended a university program. The focus is on entrepreneurs starting a new private firm in the period 2003-2005. These are roughly 80 000 individuals. Individuals born outside Sweden (roughly 18 000) are excluded, since a main concern in the current setup is whether the region of studies will exert a stronger pull on entrepreneurs than their region of birth, an analysis that is not possible for immigrant entrepreneurs. Lastly, individuals born before 1961 (roughly 28 000) are excluded since the register from which we collected this particular data lacks detailed information on the location of birth and has weaker coverage of education data for those older cohorts. The resulting sample consists of 35 187 individuals. Excluding cohorts born before 1961 is not considered a major drawback since any pull effect the region of studies might exert on an entrepreneur is expected to fade away so many years after graduation.

In the regional dimension, it is necessary to divide Sweden in a set of regions that will constitute a relevant set of alternatives from which an entrepreneur can choose when setting up his or her firm. The basis of the applied regional breakdown is the division of Sweden into 81 functional regions, constructed from labour commuting statistics of 2003[2]. For the purpose of this study, those functional regions were aggregated into 12 regions[3]. Both sets of regions are depicted in Figure 1. The main principle in this process has been to merge more sparsely populated functional regions that lack major centres of higher education to the geographically adjacent university centre, based on student registration

statistics from the Swedish national agency for higher education. A secondary principle has been to preserve the status of the three major urban functional regions of Stockholm-Uppsala, Gothenburg, and Malmö-Lund. Table 1 offers a description of these 12 regions. The dominant role of the Stockholm-Uppsala region is easily observed, exhibiting three times the number of employees and new entrepreneurs (but fewer than three times the number of students) of the other two major urban centres of Gothenburg and Malmö-Lund.

Table 1 somewhere around here

Figure 1 somewhere around here

Table 2 describes the trends in location choices apparent among the identified group of entrepreneurs. Roughly two thirds of the total population of entrepreneurs established their firm in the same region where they were born (see group 1). Of the total, 29.6% had attended a Swedish university and while 15.4% had studied in the same region they were born, 14.2%, a total of 4 979 individuals, had studied away from home (see group 2). Group 3 consists of entrepreneurs with a university degree that studied in their region of birth, 5 429 individuals[4]. Interestingly enough 4 622 of those chose to start their firm in their region of birth while only 807 chose to set up somewhere else. In other words, roughly 85% of individuals studying at their region of birth also established their firms there. Compared to the overall tendency of starting a firm at the individual's region of

birth, shown through group 1 (67.8%), it would appear that entrepreneurs that studied in their region of birth have a greater tendency to start their firm there compared to the total. Considered next are those individuals with a university degree that studied away from home (group 4). Out of these 4 979 individuals only 1 389 returned to their region of birth to start a firm, while almost half of them (2 357 entrepreneurs) set up their firms in their region of studies. The remaining 1 233 representing 3.6% of the total chose to start a firm in a region that they neither studied in, nor were born in.

Table 2 somewhere around here

These summary statistics seem to support the authors' expectations. Not only do universities seem to attract alumni entrepreneurs to stay in their region of studies when establishing a new firm, but an entrepreneur studying in his region of birth will also be less likely to start a business somewhere else.

Econometric model

In order to formally test the three hypotheses, the location choice of the individual entrepreneur is examined conditional on a vector of independent variables. Let u_{im} stand for the utility individual *i* receives from alternative *m*. Then,

$$u_{im} = \mathbf{x}_{im}\boldsymbol{\beta} + \varepsilon_{im} \;, \quad m = 1,\ldots,J \qquad , \; J = 12 \tag{1}$$

where $\mathbf{x}_{im}$ represents a vector of alternative-specific regional characteristics. $\boldsymbol{\beta}$ is the vector of coefficients capturing the weight the average individual assigns to the elements of $\mathbf{x}_{im}$, and ε_{im} is a random, normally distributed error term. The probability of choosing alternative m from among J different alternatives is:

$$\Pr(y_i = m) = \Pr(u_{im} > u_{ij}, \forall j \neq m) =$$

$$= \Pr(\mathbf{x}_{im}\boldsymbol{\beta} + \varepsilon_{im} > \mathbf{x}_{ij}\boldsymbol{\beta} + \varepsilon_{ij}, \forall j \neq m) \ , \ j = 1, \ldots, J \qquad (2)$$

That is, the probability to choose region m is equal to the probability that the utility of setting up a firm in region m for this particular entrepreneur is larger than the utility of setting up a firm in any other region. To this probability, the functional form of the conditional logit model (CLM; McFADDEN, 1974) is assigned. In the CLM the predicted probability of observing outcome m is

$$\Pr(y_i = m \mid \mathbf{x}_{im}) = \frac{\exp(\mathbf{x}_{im}\boldsymbol{\beta})}{\sum_{j=1}^{J} \exp(\mathbf{x}_{ij}\boldsymbol{\beta})} \quad \text{for } m = 1 \text{ to J} \qquad (3)$$

The elements in $\boldsymbol{\beta}$ are calculated through maximum simulated likelihood methods. Two important shortcomings of the CLM need be addressed.

First of all the response probabilities in equation (3) do not take into account individual heterogeneity. Hence, no attributes of the entrepreneur such as age or gender can be included as predictors of locational choice in a straightforward manner. Note that in

equation (3) the regressors vary along both individuals *i* and alternatives *j*. Individual characteristics (e.g. age) that vary across *i* but are fixed across *j* are not admitable in this model since they would be perfectly collinear. These response probabilities may be compared to those of the multinomial logit model where regressors vary across individuals but not across alternatives:

$$\Pr(y_i = m \mid \mathbf{x}_i) = \frac{\exp(\mathbf{x}_i \boldsymbol{\beta}_m)}{\sum_{j=1}^{J} \exp(\mathbf{x}_i \boldsymbol{\beta}_j)} \quad \text{for m = 1 to J} \qquad (4)$$

To ensure model identification the usual restriction is that $\boldsymbol{\beta}_1 = 0$ and J-1 vectors of coefficients capture the difference between the probability of choosing alternative m and alternative 1 from changes in the regressors. The two models can be combined into what some authors call a mixed logit model, where

$$\Pr(y_i = m \mid \mathbf{x}_{im}, \mathbf{w}_i) = \frac{\exp(\mathbf{x}_{im} \boldsymbol{\beta} + \mathbf{w}_i \boldsymbol{\gamma}_m)}{\sum_{j=1}^{J} \exp(\mathbf{x}_{ij} \boldsymbol{\beta} + \mathbf{w}_i \boldsymbol{\gamma}_j)} \quad \text{for m = 1 to J} \qquad (5)$$

In (5) $\mathbf{x}_{im}$ vary over alternatives (and individuals) and $\mathbf{w}_i$ vary only over individuals. A mixed logit specification is also estimated as a robustness test but the authors choose to focus attention on the conditional logit since the output is less cumbersome and more straighforward to discuss and interpret. See CAMERON and TRIVEDI (2005) pp.490 for a more in-depth discussion of the family of multinomial models.

Second of all, and most importantly, the CLM depends on the Independence of Irrelevant Alternatives (IIA) assumption. Formally, the IIA assumption states that for individual *i* the error terms across the different choices are uncorrelated. Analytically,

$$corr(\varepsilon_{ik}, \varepsilon_{il}) = 0 \ , \ \forall k \neq l \ , \quad k, l \in (1,...,12) \tag{6}$$

What this means is that an individual's preference between two of the alternatives in the choice set is not affected by the rest of the alternatives available. This is a rather strong assumption, especially in the current setup. In order to control for both of these shortcoming the main results are compared to those from an alternative specific multinomial probit model (ASMPM; LONG and FREESE, 2006) that allows for case-specific (that control for individual heterogeneity) as well as alternative specific controls, similarly to a mixed logit model of the type described by equation (5), but most importantly also relaxes the assumption of uncorrelated error terms.

In specifying the covariates in the vector $\mathbf{x}_{im}$, the individual-specific utility of setting up in region *m* is modelled as a function of three types of factors: previous experiences, agglomeration economies and region-specific idiosyncrasy. Capturing the first factor, dummy variables are included describing whether individual *i* was *Born*, had *Studied*, had recently *Worked* or not in region *m*.[5] Note that not all individuals have recent work experience or have studied at a Swedish university. Localisation economies are captured

by a production structure specialization index (PS)[6,7] which measures the extent to which region j is specialized towards industry h:

$$PS_{hj} = \left[\frac{E_{hj}}{\sum_{h} E_{hj}} \right] / \left[\frac{\sum_{j} E_{hj}}{\sum_{h}\sum_{j} E_{hj}} \right] \quad (7)$$

where h = 1, … , 43 for each industry branch, using two-digit SNI classification

j = 1, … , 12 for each region

E = employment

For practical purposes the PS-index is standardized using the formula (PS-1)/(PS+1) to make it balanced and constrained within the interval (-1,1). This way positive values of the PS-index refer to industries whose share of employment in a particular region is greater to this industry's share in national employment, while negative values refer to the exact opposite. Urbanisation economies are captured by a dummy (*Urban*) equalling one for the three major urban centres of Stockholm-Uppsala, Gothenburg, or Malmö-Lund, and zero otherwise. Finally, a full set of dummy variables, one for each alternative to an arbitrary base region, is included to control for the heterogeneous incentive structures across regions described in the last paragraphs of section 2. The reason that these regional characteristics can be lumped together is that they do not vary across individuals but provide all entrepreneurs with the same regional-specific base utility, which does vary across the alternative choices. In the basic specification, the vector of control variables becomes:

$$\mathbf{x}'_{im} = \left[Born \quad Studied \quad Worked \quad PSindex \quad Urban \quad Alternative_2 \quad \ldots \quad \ldots \quad Alternative_J \right] \tag{8}$$

A final remark, before turning to the results of the analysis, concerns the high correlation among the binary control variables describing the relation of the entrepreneurs with each region (see table 3). Having mostly binary controls can result in biased estimates, especially when these exhibit a high degree of correlation. In such cases having a clear expectation of the direction of the bias helps mitigate the problem. In the current setup *Worked* is expected to have the strongest effect on the location choice of the entrepreneur since it reflects the most recent location of the entrepreneurs prior to starting up their own firm. Most importantly, it can be expected that the inclusion of *Worked* introduces a negative bias to the effect of *Born* and *Studied* since it is very rarely the case that an individual will choose self-employment when she first enters the job market (83.3% of the identified entrepreneurs held a job in the private sector of the economy in the last 5 years prior to starting up their own firm). The current specification may therefore underestimate the significance of having studied in a region for the probability of choosing that location for starting a firm since part of that effect will be captured by the local employment history of the individual.

Table 3 somewhere around here

RESULTS

Main Results

The results of the main econometric analysis are summarized in table 4, while tables 5 and 6 and 7 report robustness checks. In table 4 a total of six different specifications have been estimated. Regression (I) is a barebone specification controlling merely for the regions of birth and studies of the individual. The effect of both dummy variables is positive and significant, controlling for regional heterogeneity (as in all specifications). Interpreting the results of the conditional logit estimation is rather straightforward. Exponentiation of the coefficients gives the change in the ratio of the odds of choosing a particular alternative over the rest from a discrete change of a dummy variable or a small change in a continuous variable. For example, considering the estimate β_1 for the control *Born,* the odds that an entrepreneur will choose to setup her firm in her region of birth over any other region is e^{β_1}. However, since the predicted probabilities are a non-linear function of the latent linear utilities the interpretation of particularly the interaction terms is somewhat troublesome since their marginal effect is not uniquely defined but will vary with the values of all regressors (AI and NORTON, 2003; NORTON et al., 2004; WIERSEMA and BOWEN, 2009), a point further addressed below.

Regression (II) adds a control for the recent employment history of the entrepreneur. This test of the sensitivity of the results to the exclusion of *Worked* is necessary given the discussion in the last paragraph of Section 3. As expected, the coefficient on *Worked* in (II) is the largest one, and controlling for the individual's employment history reduces the

impact of both *Born* and *Studied* that nevertheless remain strong and significant. Regression (III) – the core model of the analysis – further extends the model by including the *PS Index* and *Urban* dummy, controlling for the impact of localisation and urbanisation economies, respectively. This inclusion does not significantly change the effect of the three main dummy controls. Supporting the findings of FIGUEIREDO et al. (2002), the present study finds that conditioning for regional heterogeneity and his personal ties to the different regions, the externalities of localisation and urbanisation economies still exhibit positive effects on the location choice of entrepreneurs. So far, the evidence in support of the first hypothesis is quite robust. In regression (IV) an interaction term between *Studied* and *Born* is included in a further attempt to disentangle the two effects. The effect of this interaction term in negative and significant, but its inclusion drives the magnitude of *Born* and *Studied* upwards without diminishing their significance. This indicates that the average entrepreneur is more likely to remain in his or her region of birth if (s)he also attended a university program in the same region. Even controlling for the recent employment history of the entrepreneurs, and their region of birth, they still exhibit an increased propensity of setting up in their region of studies. . All findings seem to support Hypothesis 1.

Following the introduction of the interaction term between *Studied* and *Born* it is necessary at this point to elaborate on the interpretation difficulties alluded to earlier in the text. A succinct description of the issue is that due to the non-linear relationship between the latent utilities and the predicted probabilities the marginal effect of the

interaction term is not captured by the coefficient estimate on the interaction term as in the familiar OLS setting. Instead that effect is conditional on the values of all regressors (in other words on which point on the cumulative distribution function the marginal effect is considered). To address this issue the marginal effect of the interaction term and the corresponding regressors are evaluated over different combinations of the independent variables. The effect of the interaction term between *Studied* and *Born* passed successfully this test.

Turning to the formal test of the second hypothesis, regression (V) adds an interaction term between *Studied* and the *PS Index* to the specification of regression (III). The effect of the interaction term is negative and significant. In other words, localisation economies appear more significant for the location decision of new firms in the absence of an alumni effect. This finding, which holds across different combinations of the independent variables (see Appendix A for a full set of marginal effects), provides support for Hypothesis 2.

Table 4 in somewhere around here

Finally, in regression (VI) an interaction term between *Studied* and *Urban* is added to the core model. This addition does not affect the significance of the original set of dummies controlling for the personal ties of the individuals with the different regions. The interaction term itself is negative and significant. This result indicates that when

localisation decisions are driven by an alumni effect, the impact of urbanisation economies decreases, as predicted in hypothesis 3. However, even if the alumni effect is stronger in non-urban regions, urban regions also exert a similar effect on local entrepreneurial alumni. Similarly as before, the significance and direction of the marginal effect was confirmed across different combinations of the independent variables (see Appendix A).

Robustness tests

As discussed above, the use of the conditional logit estimation introduces two kinds of problems. First, it does not allow for individual characteristics to systematically affect choices. While there is no clear theory predicting that e.g. the age or gender or the entrepreneur should affect location choices, it is nonetheless preferable to check for the robustness of the results across these characteristics. The sample is therefore split according to the sex and age (younger or older than 35) of the entrepreneur in table 5, and repeat the estimation of regressions (III), (V) and (VI) of table 4 that concern the three hypotheses. The results hold equally between males and females and older and younger entrepreneurs.[8] Turning to the problem of the IIA assumption, a simple way to test whether it is violated is to test whether the results are sensitive to excluding one alternative at a time from the choice set. All results do pass this test but CHENG and LONG (2007) suggest that such tests that are based on the estimation of a restricted choice set are unsatisfactory for applied work. The results are therefore tested against those of the alternative-specific multinomial probit model (ASMPM), which allows both the inclusion

of individual specific controls and relaxation of the assumption of uncorrelated error terms. The ASMPM estimation results, that are not presented here to save space, support the premises of the CLM results.[9]

In order to test that the results do not depend on the broad regional aggregation, the CLM analysis is also repeated for the case of the original 81 functional regions. The results are presented in table 6. The direction and sign of all effects match exactly those in table 4 for the case of the 12 large regions.

Since only a third (29.6%) of the identified entrepreneurs had studied in a Swedish university, the coefficient estimates obtained for *Born* and *Worked* could be artificially high, driven by those observations for which no university track record exists. Testing the strength of this potential bias, all six specifications presented in Table 4 are re-estimated for the 10,408 entrepreneurs that attended a Swedish university and presented in Table 7. As expected, the effect of *Born* and *Worked* appear weaker when considering university graduates alone. The same does not hold for the agglomeration measures, with the coefficient on *Urban* being markedly higher in this case. The effect of *Studied* appears the same as in the full sample. It must be note that this estimation should be considered as a further robustness check rather than grounds for strict comparisons. HOETKER (2007) warns that "for cross-group differences in logit coefficients to be meaningful, each group must have the same amount of unobserved variation, that is, the variation in outcomes beyond that explained by the independent variables". What is important is that when re-estimating the core models of the analysis for university graduates alone no switch in significance or reversal of sign is

observed for any of the regressors. Furthermore, we test for the sensitivity of the first hypothesis across eight different types of education. The results, which are available upon request, show that the alumni effect is significant for all categories.

While not reported in any of the tables below, we also performed robustness tests against possible multicolinearity problems by ensuring that the sign and significance of the three correlated variables *Born*, *Studied*, and *Worked* remain constant no matter which combinations of the three variables that are included in the regression. Finally, since the paper has treated the location choice of entrepreneurs without explicitly addressing the decision to become an entrepreneur it is necessary to acknowledge the possibility that the present analysis might suffer from a particular form of selection bias. An alternative interpretation of the results could be that individuals with closer ties to their place of studies are more prone to choose self-employment as a career path, self-selecting themselves into the group of individuals considered in the present study. As a test that this is not the case the analysis is repeated on the 121,948 individuals identified in the data to switch employers between years 2004 and 2005. The focus is on people switching jobs since transcending into entrepreneurship is similarly a switch in employment (even if the switch is from unemployment). Appendix B reports the results of the CLM regression for specifications (I) and (II) of Table 4, for the case of individuals switching jobs. The signs and significances of the three spatial history dummies remain unchanged from tables 4, suggesting that the main findings of the paper do not suffer from a severe selection bias problem. Including the localisation and urbanisation controls in the analysis of the individuals switching jobs does not alter the results concerning, work, birth and studies

but since the hypotheses concerned with those two aspects of location choice were specific to entrepreneurial behaviour those extended models are not presented or discussed in the present paper.

CONCLUSIONS

In this paper, a mechanism through which universities contribute to the development of local economies is suggested and empirically tested. The point of departure is the importance of social networks in the establishment and development of a new business that has been identified in the literature of entrepreneurship. It is argued that the networks – both such that are directly related to entrepreneurial activities and networks of a purely personal nature – that entrepreneurial individuals develop in and around the universities they attend are important enough to increase their propensity to choose the region they studied in as the location for their business ventures.

This hypothesis is empirically tested by considering the location decision of 35 187 entrepreneurs born in Sweden after 1960 who founded a firm in the period 2003-2005. Controlling for regional heterogeneity, their region of birth and their recent employment history these individuals exhibit an increased likelihood of locating their firms in their region of studies providing evidence in support of our hypothesis.

The impact of universities on the localisation choices of entrepreneurs is also related to the role of agglomeration economies for these decisions. This study thus adds to an emerging research stream that is able to analyse the role of agglomeration economies in location decisions alongside with individual-level factors (FIGUEIREDO et al., 2002; DAHL and SORENSEN, 2009). On examination of urbanisation economies, the effect of universities on localisation decisions is found to be stronger in non-urban regions than in urban regions. Turning to localisation economies, the pull effect of universities is found to substitute that of localization externalities. In other words, alumni entrepreneurs tend to start firms in sectors that are previously underrepresented in the region, thereby contributing to renewal of the regional economy, the generation of new entrepreneurial opportunities (AUDRETSCH and KEILBACH, 2004) and quite possibly to regional employment growth (BISHOP and GRIPAIOS, 2010). In view of recent re-interpretations of industrial agglomerations as outcomes of localised processes of heritage rather than as the outcome of generic localisation economies (BUENSTORF and KLEPPER, 2009), this finding calls for further critical examination of the role of agglomeration economies for the location choices of entrepreneurs.

The analysis of this paper highlights the role of universities as attractors of entrepreneurial activities to a region, and points out both where (non-urban regions) and how (contributing to diversification rather than specialisation) alumni entrepreneurs make the biggest difference for a regional economy. Turning to the policy implications of this analysis, it appears that the results have implications for the debate on how regions can

stimulate local entrepreneurship. DAHL and SORENSEN (2009) suggest that entrepreneurs are embedded in the regions they have the strongest social ties with and state that governments therefore need to give up the efforts to attract migrating entrepreneurs to a region and should focus on stimulating entrepreneurship in the local population. This study supports the fundamental notion that local embeddedness is a central factor for the location choices of entrepreneurs, but offers a complementary view on the potential of migration-based policies. Regional governments may find that by ensuring the attractiveness of the region to mobile students, the foundations for a boost to local entrepreneurship may in effect have been laid.

Acknowledgement - The authors are grateful for valuable comments and suggestions from Michael S. Dahl, Michael Fritsch, Tom Petersen, Kathrin Müller and other participants at a workshop in Jena, October 2009, and the ERSA conference in Jönköping, August 2010, as well as from five anonymous referees.

NOTES

1. Throughout the text, the term 'university' is used interchangeably with the term 'higher education institution'.
2. The division of Sweden into 81 functional regions is based on analysis by the Swedish Agency for Economic and Regional Growth. The following methodology was applied in forming these regions. A municipality qualifies as being "independent"

if i) no more than 20% of its working population is employed outside the municipality and if ii) no more than 7,5% of its working population commutes to the same destination municipality. All municipalities that fail to meet one or both of these requirements are considered "dependent". "Independent" municipalities form the core of a functional region and "dependent" municipalities are considered part of the functional region whose core represents the target of the majority of their out-commuters.

3. Relatively large regions are considered suitable for the analysis, given the relatively dispersed locational pattern of HEIs (see also HOARE and CORVER, 2010). This aggregation also has the advantage that it allows applying a more robust econometric methodology. As further discussed below, it is desired to compare the results with those of the alternative specific multinomial probit model. With available software, this model can only be estimated for a maximum of 20 different choices. For a very large data set such as the one used in this study, it is necessary to reduce the number of choices even further to make estimation feasible.
4. Note that since the data does not capture individuals' mobility between birth and university studies, the actual percentage of entrepreneurs who studies "away from home" (48 percent) may be inflated. However, as the official statistics from the Swedish Agency for Higher Education report that on average 44 percent of all students move between regions when starting studies, it is believed that the size of this inflation is relatively small.
5. In constructing the variable *Studied*, we use information on the most advanced tertiary education that the individual has participated in. Hence, each individual can only be registered to have studied in at most one region. In constructing the variable *Worked*,

we operationalise the notion of "recent work history" as the five-year period prior to new firm formation.

6. A specification with the other measure commonly used in the relevant literature (BEAUDRY and SCHIFFAUEROVA, 2009), the size of the industry, was also tested without altering the results.
7. For construction of the PS Index variable, data from 2005 was used.
8. All models were also re-estimated using the mixed logit specification referred to earlier in the paper, establishing that all results hold with the additional controls of age and gender included. These estimates, which show a tendency for female and younger entrepreneurs to favor urban locations, are available upon request.
9. In applying this robustness test, it was not possible to include controls for urbanisation economies, as factors that vary across locations but not across individuals cannot be included, in the ASMPM. Furthermore, the estimation of the ASMPM is very demanding in terms of computing capacity, and with the massive data used in this study, difficulties were faced in making model V converge properly. The ASMPM robustness test therefore only applies to the main hypothesis (H1).

APPENDIX A

In the context of the conditional logit model, the coefficient estimates correspond to a linearized version of the marginal effects of each factor. Note that in order to evaluate hypotheses two and three it is necessary to test how the combined impact of the variable capturing agglomeration economies (PsIndex and Urban, respectively) and the corresponding interaction term vary over different combinations of the independent variables. Furthermore, it is appropriate to compare the linearly estimated marginal effects

with marginal effects based on probability estimates when the intercept is assumed to be zero. As shown below, estimates of marginal effects support hypotheses 2 and 3 over all values of *Worked*, *Born* and *PS index*, even if the additional marginal effect from agglomeration economies is very low in the combined presence of "personal effects" (when *Worked*, *Born* and *Studied* all are set to unity).

Table A1 somewhere around here

Table A2 somewhere around here

APPENDIX B

Table B1 somewhere around here

REFERENCES

ACS, Z.J. and ARMINGTON, C. (2004) Employment growth and entrepreneurship in cities, *Regional Studies* **28**, 911-927.

AI, C. and NORTON, E.C. (2003) Interaction terms in logit and probit models. *Economics Letters* **80**,123-129.

ALDRICH, H. and ZIMMER, C. (1986) Entrepreneurship through social networks. In D. SEXTON and R. SMILER (eds.), *The Art and Science of Entrepreneurship*, 3-23, New York: Ballinger.

ANDERSSON, M. and S. KOSTER (2010) Sources of persistence in regional start-up rates – evidence from Sweden, *Journal of Economic Geography*, first published online February 4, 2010 doi:10.1093/jeg/lbp069.

ANDERSSON, Å.E., C. ANDERSTIG, HÅRSMAN, B. (1990) Knowledge and Communications Infrastructure, *Regional Science and Urban Economics* **20**, 359-376.

ARMINGTON, C. and ACS, Z.J. (2002) The determinants of regional variation in new firm formation, *Regional Studies* **36**, 33-45.

AUDRETSCH, D. and KEILBACH, M. (2004) Entrepreneurship and economic growth: an evolutionary interpretation, *Journal of Evolutionary Economics* **14**, 605-616.

BANIA, N., EBERTS, R.W. and FOGARTY, M.S. (1993) Universities and the startup of new companies: Can we generalize from Route 128 and Silicon Valley?, *The Review of Economics and Statistics* **75**(4), 761-766.

BEAUDRY, C. and SCHIFFAUEROVA, A. (2009) Who's right, Marshall or Jacobs? The localization versus urbanization debate, *Research Policy* **38**, 318-337.

BISHOP, P. and GRIPAIOS, P. (2010) Spatial externalities, relatedness and sector employment growth in Great Britain, *Regional Studies* **44**, 443-454.

BLACKWELL, M., GIBSON, D.V. and WEINBERG, D. (2002) The economic impact of educational institutions: Issues and methodology, *Economic Development Quarterly* **16**, 88-95.

BOSCHMA, R.A. and WENTING, R. (2007) The spatial evolution of the British automobile industry: does location matter? *Industrial and Corporate Change* **16**, 213-238.

BRAUNERHJELM, P. (2008) Specialization of Regions and Universities - The New Versus the Old, *Industry and Innovation* **15**(3), 253 – 275.

BUENSTORF, G. and GUENTHER, C. (2007) No place like home - Location choice and firm survival after forced relocation in the German machine tool industry. Jena Economic Research Papers # 2007-053.

BUENSTORF, G., and KLEPPER, S. (2009) Heritage and Agglomeration: The Akron Tyre Cluster Revisitedhttp://www3.interscience.wiley.com/journal/122263535/abstract - fn1, *The Economic Journal* **119**, 705-733.

CAMERON, A.C. and TRIVEDI, P.K. (2005) *Microeconometrics: Methods and Applications,* Cambridge University Press.

CHESHIRE, P. and MALECKI, E. (2004) Growth, development, and innovation: A look backward and forward, *Papers in Regional Science* **83**, 249-267.

COOPER, A.C., FOLTA, T.B., and WOO, C. (1995) Entrepreneurial information search, *Journal of Business Venturing* **10**(2), 107-120.

DAHL, M. and SORENSEN, O. (2009) The embedded entrepreneur, *European Management Review* **6**, 172–181.

DURANTON, G. and PUGA, D. (2000) Diveristy and specialisation in cities: Why, where and when does it matter? *Urban Studies* **37**, 533-555.

DURANTON, G. and PUGA, D. (2001) Nursery cities: urban diversity, process innovation and the life cycle of products, *American Economic Review* **91**, 1454-1477.

DURANTON, G. and PUGA, D. (2004) Micro-foundations of urban agglomeration economies. In HENDERSON, V. and THISSE, J.F. (Eds.) Handbook of Regional and Urban Economics, volume 4. Amsterdam: NorthHolland, 2063–2117.

DRUCKER, J and GOLDSTEIN, H. (2007) Assessing the regional economic development impacts of universities: a review of current approaches, *International Regional Science Review* **30**, 20-46.

EVANS, D.S. and LEIGHTON, L.S. (1989) Some empirical aspects of entrepreneurship, *American Economic Review* **79**, 519-535

ETZKOWITZ, H. and KLOFSTEN, M. (2005) The Innovating Region: Toward a Theory of Knowledge-Based Regional Development, *R&D Management* **35**, 243-255.

FELDMAN, M.P. (2001) The entrepreneurial even revisited: Firm formation in a regional context. *Industrial and Corporate Change*, **10**(4), 861-891.

FELSENSTEIN, D. (1995) Dealing with 'induced migration' in university impact studies, *Research in Higher Education* **36**, 457-72.

FELSENSTEIN, D. (1996) The university in the Metropolitan arena: Impacts and public policy implications, *Urban Studies* **33**, 1565-1581.

FIGUEIREDO, O., GUIMARÃES, P. and WOODWARD, D.P. (2002) Home-field advantage: location decisions of Portuguese entrepreneurs, *Journal of Urban Economics* **52**, 341-361.

FLORAX, R. (1992) *The University: A Regional Booster?* Avebury: Aldershot.

FRITSCH, M. and, MÜLLER, P. (2007) The persistence of regional new business formation-activity over time – assessing the potential of policy promotion programs, *Journal of Evolutionary Economics* **17**, 299-315.

GERTLER, M. (2010) Rules of the game: the place of institutions in regional economic change, *Regional Studies* **44**, 1-15.

GOLDSTEIN, H.A. and RENAULT, C.S. (2004) Contributions of universities to regional economic development: a quasi-experimental approach, *Regional Studies* **38**, 733-746.

GOLDSTEIN, H.A., and DRUCKER, J. (2006) The economic development impacts of universities on regions: do size and distance matter? *Economic Development Quarterly* **20**, 22-43.

GREVE, A. and SALAFF, J.W. (2003) Social networks and entrepreneurship, *Entrepreneurship Theory and Practice* **28**(1), 1-22.

GROEN, J. A. (2004) The effect of college location on migration of college-educated labor, *Journal of Econometrics* **121**, 125-142.

HANSEN, E.L. (1995) Entrepreneurial network and new organization growth, *Entrepreneurship Theory and Practic* **19**(4), 7-19.

HISRICH, R.D. and SMILOR, R.W. (1988) The university and business incubation: Technology transfer through entrepreneurial development, *The Journal of Technology Transfer* **13**, 14-19.

HOARE, A. And CORVER, M. (2010) The regional geography of new young graduate labour in the UK, *Regional Studies* **44**, 477-494.

HOETKER, G. (2007) The use of logit and probit models in strategic management research: Critical issues, *Strategic Management Journal* **28**, 331-343.

HSU, D.H., ROBERTS, E.B. and EESLEY, C.E. (2007) Entrepreneurs from technology-based universities: Evidence from MIT, *Research Policy* **36**, 768-788.

JOHANNISSON, B. (1988) Business formation – a network approach. *Scandinavian Journal of Management* **4**(3), 83-99.

JOHANSSON, B. (2004) Parsing the Menagerie of Agglomeration and Network Externalities, CESIS Working Paper No. 2.

KEEBLE, D. and WALKER, S. (1994) New firms, small firms and dead firms: spatial patterns and determinants in the Unites Kingdom, *Regional Studies* **28**, 411-427.

KLEPPER, S. (2007) Disagreements, spinoffs, and the evolution of Detroit as the capital of the U.S. automobile industry, *Management Science* **53**, 616-631.

KRUGMAN, P. (1991) Increasing returns and economic geography. *Journal of Political Economy* **99**(3), 483-499.

LEE, S.Y., FLORIDA, R. and ACS, Z.J. (2004) Creativity and entrepreneurship: a regional analysis of new firm formation, *Regional Studies* **38**, 879-891.

LINDHOLM-DAHLSTRAND, A. (1997) Growth and inventiveness in technology-based spinoff firms, *Research Policy* **26**(3), 331-344.

LINK, A.N. and SCOTT, J.T. (2007) The economics of university research parks, *Oxford Review of Economic Policy* **23**, 661-674.

LONG, J. S., and FREESE, J. (2006) *Regression Models for Categorical Dependent Variables Using Stata*, Second Edition, Stata Press.

MARSHALL, A. (1920) *The Principles of Economics*, 8th ed., Macmillan, London.

MCFADDEN, D.L. (1974) Conditional logit analysis of qualitative choice behaviour, in ZAREMBKA, P. (Ed.) *Frontiers in Econometrics*, pp. 105-142, Academic Press, New York

MYRDAL, G. (1957) *Economic Theory and Underdeveloped Regions*. Duckworth, London.

NORTON, E.C., WANG, H. and AI, C. (2004) Computing interaction effects and standard errors in logit and probit models. *The Stata Journal* **4**, 154-167.

OHLIN, B.G. (1933) *Interregional and international trade*, Harvard University Press.

PHELPS, R. (1998) The effect of university host community size on state growth. *Economics of Education Review* **17**, 149-58.

ROSENTHAL, S.S. and STRANGE, W.C. (2001) The determinants of agglomeration, *Journal of Urban Economics* **50**, 191-229.

ROTHARMEL, F.T., AGUNG, S.D. and JIANG, L. (2007) University entrepreneurship: a taxonomy of the literature, *Industrial and Corporate Change* **16**, 691–791.

SAXENIAN, A., and HSU, J.H. (2001) The Silicon Valley-Hsinchu connection: Technical communities and industrial upgrading, *Industrial and Corporate Change* **10**(4), 893-920.

SIEGFRIED, J.J., SANDERSON, A.R. and MCHENREY, P. (2007) The economic impact of colleges and universities, *Economics of Educations Review* **26**, 546-558.

SORENSEN, O. and AUDIA, P.G. (2000) The social structure of entrepreneurial activity: geographic concentration of footwear production in the United States, 1940–1989, *The American Journal of Sociology* **106**, 424-462.

STAM, E. (2007) Why butterflies don't leave. Locational behavior of entrepreneurial firms, *Economic Geography* **83**(1), 27-50.

STERNBERG, R. and WENNEKERS, S. (2007) Determinants and effects of new business creation using global entrepreneurship monitor data, *Small Business Economics,* **24**, 193-203

STUART, T. and SORENSEN, O. (2003) The geography of opportunity: spatial heterogeneity in founding rates and the performance of biotechnology firms, *Research Policy* **32**, 229-253.

VARGA, A. (1998) *University research and regional innovation. A spatial econometric analysis of academic technology transfers*. Kluwer academic publishers; Norwell, Massachusetts, USA.

WAGNER, J. and STERNBERG, R. (2004) Start-up activities, individual characteristics, and the regional milieu: Lessons for entrepreneurship support policies from German micro data, *Annals of Regional Science* **38**, 219-240.

WEBBER, M.J. (1972) Impact of Uncertainty on Location, MIT Press, Cambridge, MA.

WIERSEMA, M.F. and BOWEN, H.P. (2009) The use of limited dependent variable techniques in strategy research: Issues and methods. *Strategic Management Journal* **30**, 679-692.

TABLES

Table 1. Characteristics of regions.

Region	Size[1]	Working population[2]	Number of new firms[3]	Number of students[4]	Type
1. Stockholm-Uppsala	14,8	690	12 623	38 379	Urban
2. Gothenburg	7,3	301	4 360	16 293	Urban
3. Malmö-Lund	7,3	253	4 090	17 391	Urban
4. East Gothia	9,9	105	1 231	9 046	Non-urban
5. West Bothnia	55,2	59	777	8 381	Non-urban
6. North Bothnia	98,2	54	715	3 766	Non-urban
7. Blekinge-East Scandia	6,1	74	904	1 571	Non-urban
8. West Mälarvalley	16,0	169	2 062	8 278	Non-urban
9. West Gothia	20,9	201	2 284	8 769	Non-urban

10. N. Sveal.-S. Norrl.	67,1	220	2 579	9 009	Non-urban
11. Småland-Gotland	35,3	230	2 327	7 533	Non-urban
12. Mid-Sweden	71,0	89	1 235	3 796	Non-urban

Source: Statistics Sweden and the Swedish National Agency for Higher Education

(1) In thousands of m²

(2) In thousands, 2003-2005 average

(3) 2003-2005 average

(4) Full time equivalent, 2005 values

Table 2. A scrutiny of the group of entrepreneurs.

Description of group	Number	Percentage of Total
(1) Total number of entrepreneurs	35 187	100
- startup at region of birth	23 870	67.8
- startup away from home	11 317	32.2
(2) Entrepreneurs that studied in a Swedish University	10 408	29.6
- studied at region of birth	5 429	15.4
- studied away from region of birth	4 979	14.2
(3) Entrepreneurs that studied at their place of birth	5 429	15.4
- startup at region of birth	4 622	13.1
- startup away from region of birth	807	2.3
(4) Entrepreneurs that studied away from place of birth	4 979	14.2
- startup at region of birth	1 389	3.9
- startup at region of studies	2 357	6.7
- startup elsewhere	1 233	3.6

Note: In each section of the table a group of entrepreneurs identified by a number from 1 to 4 is broken down in two or three subgroups identified by the dashed lines

Table 3. Pairwise correlation between regions in entrepreneur's history

	Region of startup	Region of birth	Region of studies	Region of work '03	Region of work '04	Region of work '05
Region of startup	1.00					
Region of birth	0.47	1.00				
Region of studies	0.58	0.50	1.00			
Region of work '03	0.84	0.42	0.55	1.00		
Region of work '02	0.86	0.43	0.56	0.94	1.00	
Region of work '01	0.83	0.42	0.56	0.91	0.95	1.00

Note: Applicable only for the 10 408 entrepreneurs that attended a university in Sweden
VIF Region of birth: 1.53; VIF Region of studies: 1.10; VIF Region of work: 1.53.

Table 4. Regression results of the CLM regression on the probability of choosing region m ∈ (1,12)

Variable	(I)	(II)	(III)	(IV)	(V)	(VI)
Studied	2.42***	1.76***	1.75***	2.38***	1.73***	1.95***
	[0.03]	[0.04]	[0.04]	[0.05]	[0.04]	[0.06]
Born	2.94***	2.09***	2.10***	2.27***	2.10***	2.09***
	[0.01]	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]
Worked	-	3.80***	3.78***	3.77***	3.79***	3.78***
		[0.02]	[0.02]	[0.02]	[0.02]	[0.02]
Studied x Born	-	-	-	-1.58***	-	-
				[0.08]		
PS index	-	-	0.76***	0.76***	0.85***	0.80***
			[0.07]	[0.07]	[0.07]	[0.07]
Studied x PS index	-	-	-	-	-0.90***	-
					[0.24]	
Urban	-	-	0.89***	0.87***	0.91***	0.93***
			[0.07]	[0.07]	[0.07]	[0.07]
Studied x Urban	-	-	-	-	-	-0.37***
						[0.08]

Additional controls	Full set of regional dummies	Full set of regional dummies	Full set of regional dummies	Full set of regional dummies	Full set of regional dummies	Full set of regional dummies
Pseudo R^2	0.54	0.75	0.75	0.76	0.75	0.75
Log-likelihood	-40489.56	-21691.21	-21632.64	-21423.90	-21625.51	-21621.94

Number of cases: 35 187; Number of observations: 422 244 (=35 187 x 12)
*Standard errors in brackets. Significance levels: ***: 1%*

Table 5. Regression results of the CLM regression on the probability of choosing region m∈ (1,12), comparing males to females

	Panel A						Panel B					
	Males(i)			Females(ii)			Age ≤ 35(iii)			Age > 35(ii)		
Variable	(III)	(V)	(VI)	(III)	(V)	(VI)	(III)	(V)	(VI)	(III)	(V)	(VI)
Studied	1.72***	1.78***	1.89***	1.79***	1.78***	2.06***	1.83***	1.82***	1.96***	1.61***	1.61***	1.92***
	[0.05]	[0.05]	[0.07]	[0.06]	[0.06]	[0.10]	[0.05]	[0.05]	[0.07]	[0.06]	[0.06]	[0.10]
Born	2.11***	2.11***	2.11***	2.09***	2.08***	2.08***	2.11***	2.11***	2.11***	2.14***	2.14***	2.14***
	[0.02]	[0.02]	[0.02]	[0.04]	[0.04]	[0.03]	[0.02]	[0.02]	[0.02]	[0.03]	[0.03]	[0.03]
Worked	3.76***	3.76***	3.76***	3.88***	3.88***	3.88***	3.41***	3.41***	3.41***	4.29***	4.29***	4.29***
	[0.03]	[0.03]	[0.03]	[0.05]	[0.05]	[0.05]	[0.03]	[0.03]	[0.03]	[0.04]	[0.04]	[0.04]
PS index	0.71***	0.78***	0.75***	0.93***	1.08***	0.97***	0.95***	1.03***	0.98***	0.41***	0.52***	0.46***
	[0.08]	[0.09]	[0.08]	[0.13]	[0.14]	[0.13]	[0.09]	[0.09]	[0.09]	[0.11]	[0.09]	[0.11]
Studied x PS	-	-0.76***	-	-	-1.22***	-	-	-0.71***	-	-	-1.15***	-
index		[0.28]			[0.42]			[0.28]			[0.31]	
Urban	0.85***	0.85***	0.88***	0.99***	0.99***	1.04***	0.92***	0.93***	0.95***	0.84***	0.85***	0.89***
	[0.08]	[0.08]	[0.08]	[0.12]	[0.12]	[0.12]	[0.08]	[0.08]	[0.09]	[0.11]	[0.11]	[0.11]
Studied x	-	-	-0.31***	-	-	-0.48***	-	-	-0.23**	-	-	-0.54***
Urban			[0.09]			[0.13]			[0.12]			[0.13]
Additional controls	Full set of regional dummies			Full set of regional dummies			Full set of regional dummies			Full set of regional dummies		
Pseudo R^2	0.77	0.77	0.77	0.71	0.71	0.71	0.73	0.73	0.73	0.79	0.79	0.79
Log-likelihood	-15106	-15103	-21632	-6504	-6500	-6498	-13301	-13298	-13298	-8143	-8140	-8136

(i) Number of cases: 26 092; Number of observations: 313 104 (=26 092 x 12)

(ii) Number of cases: 9 095; Number of observations: 109 140 (=9 095 x 12)
*Standard errors in brackets. Significance levels: ***: 1%*

Table 6. Regression results of the CLM regression on the probability of choosing region n∈ (1,81)

Variable	(I)	(II)	(III)	(IV)	(V)	(VI)
Studied	3.15***	2.25***	2.24***	2.99***	2.21***	2.73***
	[0.02]	[0.03]	[0.03]	[0.03]	[0.03]	[0.04]
Born	3.03***	2.24***	2.25***	2.81***	2.25***	2.24***
	[0.02]	[0.02]	[0.02]	[0.03]	[0.02]	[0.02]
Worked	-	4.50***	4.48***	4.42***	4.48***	4.44***
		[0.02]	[0.02]	[0.02]	[0.02]	[0.02]
Studied x Born	-	-	-	-1.85***	-	-
				[0.06]		
PS index	-	-	0.64***	0.63***	0.82***	0.65***
			[0.05]	[0.05]	[0.05]	[0.05]
Studied x PS index	-	-	-	-	-1.33***	-
					[0.13]	
Urban	-	-	2.07***	1.98***	2.08***	2.365***
			[0.16]	[0.16]	[0.16]	[0.16]
Studied x Urban	-	-	-	-	-	-1.14***
						[0.06]
Additional	Full set of	Full set of	Full set of	Full set of	Full set of	Full set of

controls	regional dummies	regional dummies	regional dummies	regional dummies	regional dummies	regional dummies
Pseudo R^2	0.64	0.78	0.79	0.79	0.79	0.78
Log-likelihood	-58383.07	-34543.82	-34436.41	-33960.52	-34386.20	-34221.51

Number of cases: 35 187; Number of observations: 2 850 147 (=35 187 x 81)
*Standard errors in brackets. Significance levels: ***: 1%*

Table 7. Regression results of the CLM regression on the probability of choosing region m∈ (1,12) for 10 408 university graduates

Variable	(I)	(II)	(III)	(IV)	(V)	(VI)
Studied	2.21*** [0.03]	1.76*** [0.03]	1.77*** [0.04]	2.05*** [0.05]	1.74*** [0.03]	2.14*** [0.06]
Born	1.78*** [0.03]	1.37*** [0.04]	1.37*** [0.04]	1.70*** [0.05]	1.38*** [0.02]	1.38*** [0.03]
Worked	-	3.02*** [0.04]	3.01*** [0.04]	2.99*** [0.04]	3.00*** [0.04]	2.99*** [0.04]
Studied x Born	-	-	-	-0.83*** [0.08]	-	-
PS index	-	-	0.79*** [0.13]	0.78*** [0.07]	1.34*** [0.16]	0.77*** [0.13]
Studied x PS index	-	-	-	-	-1.64*** [0.25]	-
Urban	-	-	1.14*** [0.11]	1.11*** [0.07]	1.15*** [0.11]	1.38*** [0.11]
Studied x Urban	-	-	-	-	-	-0.70*** [0.08]
Additional controls	Full set of regional dummies	Full set of regional dummies	Full set of regional dummies	Full set of regional dummies	Full set of regional dummies	Full set of regional dummies
Pseudo R^2	0.58	0.71	0.71	0.71	0.71	0.71

Log-likelihood	-10910.42	-7605.85	-7587.90	-7551.21	-7566.11	-7551.59

Number of cases: 10 408; Number of observations: 124 896 (=10 408 x 12)
*Standard errors in brackets. Significance levels: ***: 1%*

Table A1. Marginal effects following estimation (V) for *PS index* and interaction term *Studied* x *PS index*

Estimation at	PS index	Studied x PS index
1. Studied = 1, rest at their median values	0.606*** [0.006]	-0.638*** [0.017]
2. Studied = 0, rest at their median values	0.167*** [0.015]	-0.176*** [0.047]
3. Born=1, Worked=1, Studied=1, rest at their median values	0.000*** [0.000]	-0.000*** [0.000]
4. Rest at their median values		
i. PS index = -0.9	0.185*** [0.16]	-0.195*** [0.051]
ii. PS index = -0.4	0.179*** [0.016]	-0.188*** [0.050]
iii. PS index = 0.1	0.160*** [0.014]	-0.168*** [0.044]
iv. PS index = 0.5	0.139*** [0.010]	-0.146*** [0.038]

Table A2. Marginal effects following estimation (VI) for *Urban* and interaction term *Studied* x *Urban*

Estimation at	Urban	Studied x Urban
1. Studied = 1, rest at their median values	0.055*** [0.032]	-0.021*** [0.004]
2. Studied = 0, rest at their median values	0.182*** [0.010]	-0.072*** [0.014]
3. Studied=1		
i. PS index = -0.9	0.094*** [0.006]	-0.037*** [0.007]
ii. PS index = -0.4	0.068*** [0.004]	-0.027*** [0.005]
iii. PS index = 0.1	0.049*** [0.002]	-0.019*** [0.003]
iv. PS index = 0.5	0.038*** [0.003]	-0.015*** [0.003]
4. Worked=1, Born=1		
i. PS index = -0.9	0.003*** [0.002]	-0.001*** [0.000]
ii. PS index = -0.4	0.002***	-0.001***

	[0.000]	[0.000]
iii. PS index = 0.1	0.001***	-0.001***
	[0.000]	[0.000]
iv. PS index = 0.5	0.001***	-0.000***
	[0.000]	[0.000]

Table B1. Regression results of the CLM regression on the probability of choosing region m ∈ (1,12) for individuals switching jobs

Variable	(I)	(II)
Studied	1.12***	0.56***
	[0.03]	[0.03]
Born	1.68***	0.63***
	[0.01]	[0.01]
Worked	-	1.754***
		[0.01]
Additional controls	Full set of regional dummies	Full set of regional dummies
Pseudo R^2	0.17	0.20
Log-likelihood	-251657.38	-242874.77

Number of cases: 121 948; Number of observations: 1 463 376 (=121 948 x 12)
Standard errors in brackets. Significance levels: ***: 1%

Figure 1. The 81 functional regions of Sweden and an aggregated set of 12 regions.

Legend: Numbers refer to the numbers assigned to regions in Table 1.